AD-A044 911 STANFORD UNIV CALIF DEPT OF OPERATIONS RESEARCH F/G 12/2

THE DIFFUSION APPROXIMATION FOR TANDEM QUEUES IN HEAVY TRAFFIC.(U)

AUG 77 J M HARRISON N00014-76-C-0418

UNCLASSIFIED TR-75 NL

1 OF 1

AD
A044911

END

DATE FILMED

10-77

DDC

THE DIFFUSION APPROXIMATION FOR TANDEM QUEUES IN HEAVY TRAFFIC

BY

J. MICHAEL HARRISON

TECHNICAL REPORT NO. 75

AUGUST 10, 1977

PREPARED UNDER CONTRACT
N00014-76-C-0418 (NR-047-061)
FOR THE OFFICE OF NAVAL RESEARCH

DEPARTMENT OF OPERATIONS RESEARCH
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

THE DIFFUSION APPROXIMATION FOR TANDEM QUEUES IN HEAVY TRAFFIC

BY

J. MICHAEL HARRISON

TECHNICAL REPORT NO. 75

AUGUST 10, 1977

PREPARED UNDER CONTRACT

N00014-76-C-0418 (NR-047-061)

FOR THE OFFICE OF NAVAL RESEARCH

Frederick S. Hillier, Project Director

This research was supported in part by National Science Foundation Grant ENG 75-14847 Department of Operations Research, Stanford University and issued as Technical Report 45

DEPARTMENT OF OPERATIONS RESEARCH

STANFORD UNIVERSITY

STANFORD, CALIFORNIA

THE DIFFUSION APPROXIMATION FOR TANDEM QUEUES IN HEAVY TRAFFIC

J. Michael Harrison

Graduate School of Business
Stanford University

1. Introduction

We consider in this paper a two-dimensional diffusion process that arises in conjunction with the following type of queuing system. The system is composed of two single server facilities (or stations) arranged in tandem. Customers arrive individually from outside the system and queue up for service at the first station. Having completed service there, they proceed to a queue in front of the second station, and after completing service at the second station they depart the system. Service is by order of arrival at each station.

We assume that the system is empty at time $t = 0$, and that the interarrival times and service times at the two stations form three mutually independent sequences of IID random variables. The interarrival times are distributed as a random variable v_0 having mean a_0 and finite variance γ_0^2, and the service times at station k $(k = 1,2)$ are distributed as a random variable v_k with mean a_k and finite variance γ_k^2. We define the traffic intensities $\rho_k = a_k/a_0$ $(k = 1,2)$. The system is said to be stable if $\rho_k < 1$ for each station k, and a situation of heavy traffic is said to prevail if ρ_k is near one for each station. For a stable system we further define

$$\alpha = (a_0 - a_1) \wedge (a_0 - a_2) = a_0[(1-\rho_1) \wedge (1-\rho_2)] \quad .$$

Thus a stable system is in heavy traffic when α is positive but near zero. Finally, it will be convenient to define

$$\mu = \begin{bmatrix} \mu_1 \\ \\ \mu_2 \end{bmatrix} \qquad \text{and} \qquad \Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12}^2 \\ \\ \sigma_{12}^2 & \sigma_2^2 \end{bmatrix} ,$$

where $\sigma_1^2 = \gamma_0^2 + \gamma_1^2$, $\sigma_{12}^2 = -\gamma_1^2$, $\sigma_2^2 = \gamma_1^2 + \gamma_2^2$, $\mu_1 = (a_1 - a_0)/\alpha$ and $\mu_2 = (a_2 - a_1)/\alpha$. Note that $\alpha\mu$ and Σ are the mean vector and covariance matrix respectively of the random vector $(v_1 - v_0,\ v_2 - v_1)$. Because we do not require that the variances γ_k^2 be strictly positive, Σ need not be positive definite, but it is non-negative definite in any case.

Let W_n^k be the waiting time (exclusive of service time) at station k for the n^{th} arriving customer, and let $W_n = (W_n^1, W_n^2)$. If the system is stable, it is well known that there exists a random vector $W = (W^1, W^2)$ such that $W_n \Longrightarrow W$ as $n \to \infty$, where $\Longrightarrow$ denotes weak convergence (convergence in distribution).

In an earlier paper [2], a limit theorem was proved to show that in heavy traffic the distribution of αW can be approximated by the limit distribution of a certain stochastic process $Z = \{Z(t),\ t \geq 0\}$. The process Z was defined as an explicit, but relatively complicated, transformation of a three-dimensional Brownian Motion. Its limit distribution was determined for one special case (in our current notation, the case $\gamma_0^2 = \gamma_1^2 = \gamma_2^2$) by an invariance principle argument, but the general problem of evaluating the limit distribution was left open. It is to this problem that we now return. The current discussion is restricted to the case of two queues in tandem, but most of what will be said extends readily to the general model with K stations in series.

In Sections 2 and 3 we repeat the definition of Z and show that it is a diffusion process (continuous strong Markov process) whose state space is the non-negative quadrant. More specifically, on the interior of its state space the process behaves like ordinary Brownian Motion with drift vector μ and covariance matrix Σ, and it reflects instantaneously at each axis (boundary surface). At one axis (the one corresponding to an empty second station in the original queueing system), the direction of reflection is normal to the axis, but at the other axis the reflection has a tangential component as well. In Section 3 we also display the (weak infinitesimal) generator of Z.

In Sections 4 through 6 we demonstrate that the limit distribution of Z is the solution of a first passage problem for a certain dual diffusion process Z^*. Using the analytical theory of Markov processes, we then derive in Section 7 the following partial differential equation (with boundary conditions) for the density f of the limit distribution.

$$[\tfrac{1}{2}\sigma_1^2 f_{11} + \sigma_{12}^2 f_{12} + \tfrac{1}{2}\sigma_2^2 f_{22} - \mu_1 f_1 - \mu_2 f_2](x,y) = 0 ,$$

$$[\sigma_{12}^2 f_{12} + \tfrac{1}{2}\sigma_2^2 f_2 - \mu_2 f](x,0) = 0 ,$$

$$[\tfrac{1}{2}\sigma_1^2 f_1 + (\tfrac{1}{2}\sigma_1^2 + \sigma_{12}^2) f_2 - \mu_1 f](0,y) = 0 .$$

(Here, and later in the paper, we use the notation

$$f_1(x,y) = \frac{\partial}{\partial x} f(x,y) , \qquad f_{12}(x,y) = \frac{\partial^2}{\partial x\, \partial y} f(x,y) ,$$

and so forth.) In Section 8 we show that the limit distribution has independent components iff $\sigma_1^2 + 2\sigma_{12}^2 = 0$, in which case each component is exponentially distributed. This expands slightly the class of solutions obtained previously in [2]. Finally, in Section 9, we explicitly solve for f when Σ is the identity matrix and $\mu_2 = 0$. (In terms of the original tandem queuing system, this corresponds to the case $\gamma_0^2 = \gamma_2^2$, deterministic services at the first station, and $\rho_1 = \rho_2$.) The distribution does not have independent components.

The explicit solutions obtained here are certainly fragmentary, but we hope that they can help to stimulate and direct the interest of others in this problem. To determine explicitly the general solution of the partial differential equation displayed above, one must overcome some significant analytical difficulties. Relatively little effort has been devoted to the problem thus far, however, and an explicit general solution may in fact be obtainable.

Finally, we note that tandem queues are the simplest case of a queuing network, so any analytical insights obtained here should be valuable in determining heavy traffic approximations for the general network model. Limit theorems justifying diffusion approximations for general networks will be presented in a forthcoming paper by Reiman [5], together with analytical characterizations of the relevant diffusions.

2. A Reflection Mapping in Two-Dimensions

Let B denote the set of functions $x(\cdot) = (x_1(\cdot), x_2(\cdot))$ that map $[0,\infty)$ into $\mathbb{R}^2$ and are bounded on finite intervals. Let B_+ be the subset of such functions that are non-negative and non-decreasing. For $x \in B$ let $C(x)$ be the set of all $\xi \in B_+$ such that

$$x_1(t) + \xi_1(t) \geq 0 \qquad \text{for all } t \geq 0$$

and

$$x_2(t) - \xi_1(t) + \xi_2(t) \geq 0 \qquad \text{for all } t \geq 0 .$$

It is obvious that $C(x)$ is non-empty. Moreover, it has a minimal element, meaning that there exists $y \in C(x)$ such that $y_1(t) \leq \xi_1(t)$ and $y_2(t) \leq \xi_2(t)$ for all $\xi \in C(x)$ and all $t \geq 0$. This minimal element is explicitly given by

$$\text{(1)} \qquad \begin{aligned} y_1(t) &= \sup_{0 \leq u \leq t} \{-x_1(u)\}^+ \\ y_2(t) &= \sup_{0 \leq u \leq t} \{-[x_2(u) - y_1(u)]\}^+ = \sup_{0 \leq v \leq u \leq t} \{[-x_1(v)]^+ - x_2(u)\}^+ . \end{aligned}$$

We define a mapping $f = (f_1, f_2): B \to B_+$ by setting $y_1 = f_1(x)$ and $y_2 = f_2(x)$. Also, we define $g = (g_1, g_2): B \to B$ by setting $g_1(x) = x_1 + y_1$ and $g_2(x) = x_2 - y_1 + y_2$.

An interpretation of the mappings f and g is as follows. Imagine a man (hereafter called the controller) who monitors two bank accounts, the contents of which fluctuate due to uncontrollable deposits and withdrawals. In the absence any intervention by the controller, the content of account k at time t will be $x_k(t)$. In particular, the initial contents of the two accounts are $x_1(0)$ and $x_2(0)$. The controller has the ability to transfer money from account 2 to account 1, and we denote by $\xi_1(t)$ the total amount of money so transferred during the interval $[0,t]$. He also can deposit money (gotten outside the system) in account 2, and we denote by $\xi_2(t)$ the total amount deposited during $[0,t]$. The pair $\xi = (\xi_1, \xi_2)$ is called a control. If we suppose that the controller is required to keep the content of each account non-negative, then $C(x)$ represents the set of feasible controls. If we suppose that the controller wants to minimize the total volume of his transactions (in the strongest possible sense), then $y = f(x)$ represents the optimal control, and the content of account k at time t will be $z_k(t)$ when the optimal control is employed, where $z = g(x)$.

The effect of the reflection mapping g is portrayed graphically in Figure 1 below. Suppose $x \in B$ is continuous and represents the sample path of a stochastic process. Let $y = f(x)$ and $z = g(x)$. Continuing the interpretation of f and g given above, one might call x the uncontrolled process and z the controlled process. If $x(0)$ is strictly positive, then $x = z$ until the first time that x strikes a boundary point of the non-negative quadrant. If the horizontal axis $(x_2 = 0)$ is struck first, then the local behavior of the controlled process after hitting is given by $z_1 = x_1$ and $z_2 = x_2 + y_2$. The control y_2 increases in the minimal amounts necessary to keep z_2 non-negative, and we have the normal reflection pictured in Figure 1. If the vertical axis $(x_1 = 0)$

Figure 1

Angles of Reflection Induced by the Mapping g

is struck first, however, then the effect of the mapping g is slightly more complicated. The local behavior of the controlled process after hitting is now given by $z_1 = x_1 + y_1$ and $z_2 = x_2 - y_1$. The control y_1 increases in the minimal amounts necessary to keep z_1 non-negative, and we have non-normal reflection in the direction shown in Figure 1.

A final important property of the mapping g is that it is memoryless in the following sense. Suppose $x \in B$ and $z = g(x)$. For arbitrary $T > 0$, let $z_T(t) = z(T + t)$ and

$$x_T(t) = z(T) + [x(T + t) - x(T)]$$

for $t \geq 0$. Then

$$z_T(t) = g(X_T)(t) \qquad \text{for } t \geq 0 . \tag{2}$$

Thus the sample path of the controlled or reflected process z after time T is completely determined by $z(T)$ and the sample path of the uncontrolled process x after time T. To be specific, it is gotten by applying the reflection mapping g to the process X_T. The memoryless property (2) follows immediately from the characterization of $f(x)$ as the minimal element of $C(x)$ but is not at all apparent from the explicit representation (1).

3. The Diffusion

Viewing μ and Σ as arbitrary data hereafter (except for the requirement that Σ be symmetric and non-negative definite), let $X = \{X(t),\ t \geq 0\}$ be a two-dimensional Brownian Motion with drift vector μ, covariance matrix Σ and general starting state. If $\sigma_1^2 = \gamma_0^2 + \gamma_1^2$, $\sigma_{12}^2 = -\gamma_1^2$ and $\sigma_2^2 = \gamma_1^2 + \gamma_2^2$ (see section 1), then such a process can be constructed by taking

$$X(t) = \begin{bmatrix} X_1(t) \\ X_2(t) \end{bmatrix} = \begin{bmatrix} \gamma_0 \xi_0(t) - \gamma_1 \xi_1(t) + \mu_1 t \\ \gamma_1 \xi_1(t) - \gamma_2 \xi_2(t) + \Sigma_2 t \end{bmatrix} , \qquad t \geq 0 ,$$

where the processes ξ_k are independent standard (zero mean and unit variance) Brownian Motions with general starting states. In the usual way, we denote by $P_x(\cdot)$ the distribution on the path space of X corresponding to starting state $x = (x_1, x_2) \in \mathbb{R}^2$. We define the processes $Y = f(X)$ and $Z = g(X)$, so that

$$Z_1(t) = X_1(t) + Y_1(t)\ , \qquad t \geq 0\ ,$$

$$Z_2(t) = X_2(t) - Y_1(t) + Y_2(t)\ , \qquad t \geq 0\ .$$

<u>Proposition 1</u>. Z is a strong Markov process with state space $S = \mathbb{R}_+^2$, and its sample paths are almost surely continuous.

<u>Proof</u>. From the explicit representation for $y = f(x)$ that was given in Section 2 it follows easily that f and g both map continuous functions into continuous functions. Thus the path continuity of Z follows from the continuity of X. The strong Markov property is immediate from (2) and the strong Markov property and stationary, independent increments of .

Observe that $Z(0) =$ (0) whenever $X(0) \in S$. Thus $P_x(\cdot)$ can alternately be interpreted as the distribution on the path space of Z corresponding to $Z(0) = x$, provided that $x \in S$. In the interest of characterizing more completely our underlying diffusion Z, we conclude this section with a result showing how its generator operates on a class of smooth functions. We shall make no further use of the generator of Z in this paper, however.

Following Dynkin [1], we define $\mathcal{L}$ to be the set of bounded, measurable functions $f: S \to \mathbb{R}$ such that $T_t f \to f$ as $t \downarrow 0$, where $T_t f(x) = E_x[f(Z(t))]$. We denote by $\mathcal{D}$ the set of $f \in \mathcal{L}$ such that $(T_t f - f)/t$ converges boundedly pointwise as $t \downarrow 0$ to a limit in $\mathcal{L}$.

The limit is denoted Gf, the operator G thus defined is called the weak infinitesimal generator of Z (or just generator for short), and $\mathcal{D}$ is called the domain of the generator.

Let $\mathcal{C}$ denote the set of bounded, continuous functions $f:S \to \mathbb{R}$ such that all partial derivatives of f exist and are bounded and continuous on S. Let $\mathcal{S}$ (for smooth) be the set of functions $f \in \mathcal{C}$ that vanish on some neighbohood of the origin. In saying that f has continuous partial derivatives on the boundary of S, we mean that each partial derivative approaches a finite limit at the boundary and the limit is a continuous function of the boundary point. Symbols like $f_2(x_1,0)$ are understood throughout to be shorthand notations for such limits.

From the path continuity of Z and bounded convergence it follows that every bounded, continuous $f:S \to \mathbb{R}$ is in $\mathcal{L}$. Thus $\mathcal{S} \subset \mathcal{L}$. For ease of notation, we define a differential operator D by

$$Df = \frac{1}{2}\sigma_1^2 f_{11} + \sigma_{12}^2 f_{12} + \frac{1}{2}\sigma_2^2 f_{22} + \mu_1 f_1 + \mu_2 f_2 .$$

The following proposition will not be proved, since it will be subsumed by a more general result of Reiman's [5] and it is not used in the remainder of this paper. A very similar result, Proposition 6 of Section 5, will be proved in detail, however.

<u>Proposition 2</u>. Suppose $f \in \mathcal{S}$. Then $f \in \mathcal{D}$ iff

(3) $$f_2(x_1,0) = 0 \quad \text{for} \quad x_1 > 0,$$

and

(4) $$(f_1 - f_2)(0,x_2) = 0 \quad \text{for} \quad x_2 > 0,$$

in which case $Gf = Df \in \mathcal{S}$.

Conditions (3) and (4) require that at each boundary surface (axis) the directional derivative in the direction shown in Figure 1 be zero. In the definitional system of Watanabe [7], such restrictions on the domain of the generator are used to define direction of reflection at a boundary point. Thus, from the perspective of Watanabe's theory, Proposition 2 justifies our statement that the diffusion Z reflects instantaneously at its boundary, the angle of reflection being as shown in Figure 1.

4. The Limit Distribution

Assuming that the process Z begins at the origin, we wish now to characterize (and if possible compute) the distributions

$$F_t(z) = P_0\{Z(t) \leq z\} \quad \text{for } t \geq 0 \text{ and } z \in S .$$

It will actually be more convenient to work with

$$\Phi_t(z_1,z_2) = P_0\{Z_1(t) \leq z_1,\ Z_1(t) + Z_2(t) \leq z_1 + z_2\} ,$$

defined for $t \geq 0$ and $(z_1,z_2) \in S$. The reason is the following result, which is precisely analogous to Lemma 1 of [2].

Proposition 3. $\Phi_t(z_1,z_2) = P_0\{M_1(t) \leq z_1,\ M_2(t) \leq z_1 + z_2\}$, where

$$M_1(t) = \sup_{0 \leq u \leq t} \{X_1(u)\} ,$$

$$M_2(t) = \sup_{0 \leq v \leq u \leq t} \{X_2(v) + X_1(u)\} .$$

Proof. The argument is so similar to the proof of Lemma 1 in [2] that we shall merely sketch it. First use (1) to write down an explicit representation for $Z(t) = (Z_1(t), Z_2(t))$, observing that the general representation simplifies considerably when $X(0) = 0$. Let $t > 0$ be fixed. Assuming that $X(0) = 0$, observe that $\{X(u), 0 \leq u \leq t\}$ has the same distribution as

$$\{X(t) - X(t-u), 0 \leq u \leq t\} ,$$

since X has stationary independent increments. Making this substitution in the representation for $Z(t)$ and simplifying, we obtain the desired proposition.

With $X(0) = 0$, it is clear that $M_1(\cdot)$ and $M_2(\cdot)$ are almost surely non-decreasing, so the limits

$$M_1 = \lim_{t\to\infty} M_1(t) \qquad \text{and} \qquad M_2 = \lim_{t\to\infty} M_2(t)$$

exist almost surely (P_0). Furthermore,

$$F_t(z) \to F(z) \quad \text{and} \quad \Phi_t(z) \to \Phi(z) \qquad \text{as} \quad t \to \infty$$

for each $z \in S$, where

(5) $\qquad F(z_1, z_2) = P_0\{M_1 \leq z_1, M_2 - M_1 \leq z_2\}$

and

(6) $\qquad \Phi(z_1, z_2) = P_0\{M_1 \leq z_1, M_2 \leq z_1 + z_2\} \qquad$ for $(z_1, z_2) \in S$.

Finally, from the strong law for Brownian Motion it follows that $M_1 < \infty$ iff $\mu_1 < 0$, whereas $M_2 < \infty$ iff $\mu_1 < 0$ and $\mu_1 + \mu_2 < 0$. Thus we have the following.

Proposition 4. The limit distribution F is proper iff $\mu_1 < 0$ and $\mu_1 + \mu_2 < 0$.

By definition, M_1 is the maximum of the unrestricted Brownian Motion X_1. When $\mu_1 < 0$ and $X_1(0) = 0$, M_1 is known to have an exponential distribution with mean $\sigma_1^2/2|\mu_1|$, and this gives the first marginal distribution of F. The same result can be obtained by observing that Z_1 is just X_1 modified by a reflecting barrier at zero, and then using well known results for reflected Brownian Motion. Note that Z_2 is not just X_2 modified by a reflecting barrier at zero. In fact, the one-dimensional process Z_2 is not Markov.

5. A Related Diffusion Process

Let $X^*(t) = -X(t)$ for $t \geq 0$, and let $P_x^*(\cdot)$ be the distribution on the path space of X^* corresponding to $X^*(0) = x$ for $x \in \mathbb{R}^2$. (This is easier than writing negative subscripts.) Next let

$$Y_2^*(t) = \sup_{0 \leq u \leq t} \{-X_2^*(u)\}^+ , \qquad t \geq 0,$$

and

$$T^* = \inf\{t \geq 0 : X_1^*(t) - Y_2^*(t) \leq 0\} ,$$

with $T^* = \infty$ if the indicated t-set is empty. Now define a process $Z^* = \{Z^*(t),\ t \geq 0\}$ by setting

$$Z_1^*(t) = \begin{cases} X_1^*(t) - Y_2^*(t) & \text{if } 0 \leq t < T^* \\ X_1^*(T^*) - Y_2^*(T^*) & \text{if } t \geq T^*, \end{cases}$$

$$Z_2^*(t) = \begin{cases} X_2^*(t) + Y_2^*(t) & \text{if } 0 \leq t < T^* \\ X_2^*(T^*) + Y_2^*(T^*) & \text{if } t \geq T^*. \end{cases}$$

Observe that, with Z_2^* defined in terms of Y_2^* in this way, Y_2^* is the minimal function of X^* such that $Z_2^*(t) \geq 0$ for all $t \geq 0$. Hereafter we restrict attention to paths of X^* that begin in the non-negative quadrant, this insuring that $Z^*(\cdot) \in S$.

The process Z^* is obtained from X^* in two stages. First, we impose an instantaneously reflecting barrier at the axis $X_2^* = 0$, the angle of reflection being as shown in Figure 2 below. Second, an

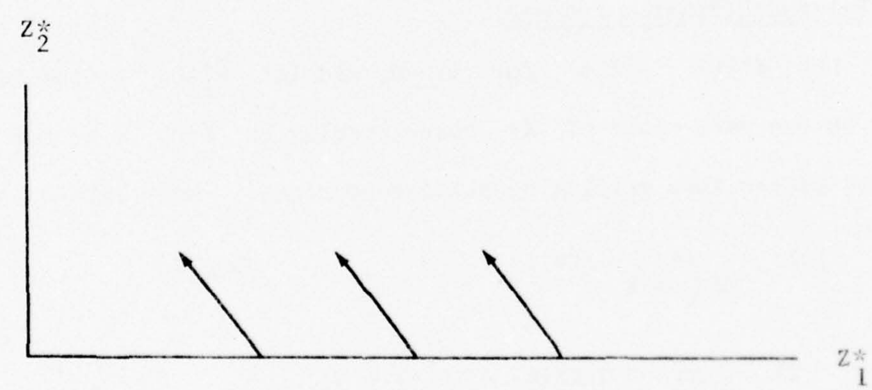

Figure 2

Angle of Reflection for the Process Z^*

absorbing barrier is imposed at the axis $Z_1^* = 0$. The proof of the following is virtually identical to the proof of Proposition 1, so we delete it.

<u>Proposition 5</u>. Z^* is a strong Markov process with state space S, and its sample paths are almost surely continuous.

Let G^* be the weak infinitesimal generator of Z^* and let $\mathcal{D}^*$ be its domain. Let $\mathcal{C}$ be defined as in Section 3. Finally, define a differential operator D^* by

$$D^*f = \frac{1}{2}\sigma_1^2 f_{11} + \sigma_{12}^2 f_{12} + \frac{1}{2}\sigma_2^2 f_{22} - \mu_1 f_1 - \mu_2 f_2 .$$

<u>Proposition 6</u>. Suppose $f \in \mathcal{C}$. Then $f \in \mathcal{D}^*$ iff $(f_2 - f_1)(x_1, 0) = 0$ for $x_1 > 0$. In this case, $G^*f(z) = 0$ for $z = (0, z_2)$ and $G^*f = D^*f$ otherwise.

<u>Proof</u>. Assume $f \in \mathcal{C}$ and let $T_t^*f(z) = E_z^*[f(Z^*(t))]$ for $z \in S$ and $t \geq 0$. We first determine the conditions under which $(T_t^*f - f)/t$ converges pointwise as $t \downarrow 0$.

For each interior point z of S, let $h_t(z)$ denote the probability that, starting in state z, the boundary is struck by Z^* before time t. Using nothing more than the first exit distribution of one-dimensional Brownian Motion (because our state space is a rectangle), it is easy to show that $h_t(z) = o(t)$ for each interior point z. Thus

$$T_t^* f(z) = E_z^*[f(X^*(t))] + o(t) ,$$

because $Z^*(t) = X^*(t)$ until the first hitting of the boundary. It follows that $[T_t^* f(z) - f(z)]/t \to D^* f(z)$ as $t \downarrow 0$, since D^* is the generator of X^* on $\mathcal{C}$. (See, for example, section 3.2 of Itô [4].)

If $z = (0, z_2)$ with $z_2 \geq 0$, then $T_t^* f(z) = f(z)$ for all $t \geq 0$, since $Z^*(t) = z$ for all $t \geq 0$. Thus $[T_t^* f(z) - f(z)]/t \to 0$ as $t \downarrow 0$. To examine the boundary $z_2 = 0$, we need some more notation. For purposes of this proof only, let

$$M(t) = \sup_{0 \leq u \leq t} [X_2^*(u) - X_2^*(0)] ,$$

$$M^*(t) = \sup_{0 \leq u \leq t} [X_2(u) - X_2(0)] = \sup_{0 \leq u \leq t} [X_2^*(0) - X_2^*(u)]$$

for $t \geq 0$. Thus $Y_2^*(t) = [M^*(t) - X_2^*(0)]^+$ for $t \geq 0$. Observe that the distributions of the Brownian maxima $M^*(t)$ and $M(t)$ do not depend on $X_2^*(0)$. Furthermore, the maximum distribution is known explicitly, and direct computation reveals that

(7) $\quad E_z^*[M^*(t)]/t \to \infty$ and $E_z^*[M(t)]/t \to \infty$ as $t \downarrow 0$,

(8) $\quad E_z^*[(M^*(t))^2]/t \to \sigma_2^2$ and $E_z^*[(M(t))^2]/t \to \sigma_2^2$ as $t \downarrow 0$

Until further notice, let $z = (z_1, 0)$ with $z_1 > 0$. It is well known that $X_2^*(t) + M^*(t)$ then has the same distribution as $M(t)$ for each

$t > 0$. (See the proof of Proposition 3 above.) Thus,

$$(9) \qquad E_z^*[(X_2^*(t))^2 + 2X_2^*(t)\ M^*(t) + (M^*(t))^2] = E_z^*[(M(t))^2] .$$

Dividing both sides of (9) by t, letting $t \downarrow 0$, and using (8), we then obtain

$$(10) \qquad E_z^*[X_2^*(t)\ M^*(t)]/t \to -\frac{1}{2}\sigma_2^2 \qquad \text{as} \qquad t \downarrow 0 .$$

Finally, letting $\Delta X_1^*(t) = X_1^*(t) - X_1^*(0)$, we need

$$(11) \qquad E_z^*[\Delta X_1^*(t)\ M^*(t)]/t \to -\frac{1}{2}\sigma_{12}^2 \qquad \text{as} \qquad t \downarrow 0 .$$

For the case $\mu_1 = \mu_2 = 0$, this is proved as follows. First observe that

$$(12) \qquad E_z^*[\Delta X_1^*(t)\ M^*(t)] = E_z^*[\Delta X_1^*(t)\ (M(t) - X_2^*(t))]$$

for $t \geq 0$. (See again the proof of Proposition 3 above.) Also, symmetry gives (when $\mu_1 = \mu_2 = 0$)

$$(13) \qquad E_z^*[-\Delta X_1^*(t)\ M^*(t)] = E_z^*[\Delta X_1^*(t)\ M(t)] .$$

Combining (12) and (13) gives (11). The general case is proved by reducing it to the driftless case through obvious bounds.

Assuming as above that $X_2^*(0) = Z_2^*(0) = 0$ and $Z_1^*(0) > 0$, we define $\Delta X_1^*(t)$ as before and

(14) $\quad \Delta Z_1^*(t) = \Delta X_1^*(t) - M^*(t), \qquad \Delta Z_2^*(t) = X_2^*(t) + M^*(t)$.

Then

(15) $\quad T_t^* f(z) = E_z^*[f(z_1 + \Delta Z_1^*(t),\ Z_2^*(t))] + o(t)$,

the $o(t)$ term on the right corresponding to the event $\{T^* < t\}$. Now an application of Taylor's Theorem (with the exact form of the remainder) gives, when combined with (15),

(16) $$T_t^* f(z) = f(z) + f_1(z)\, E_z^*[\Delta Z_1^*(t)] + f_2(z)\, E_z^*[\Delta Z_2^*(t)]$$
$$+ \frac{1}{2} f_{11}(z)\, E_z^*[(\Delta Z_1^*(t))^2] + \frac{1}{2} f_{22}(z)\, E_z^*[(\Delta Z_2^*(t))^2]$$
$$+ f_{12}(z)\, E_z^*[\Delta Z_1^*(t)\ \Delta Z_2^*(t)] + R_t(z) + o(t) ,$$

where $R_t(z)$ is a remainder term. Furthermore, since the third order partial derivatives of $f \in \mathcal{C}$ are bounded, it is easy to show that $R_t(z) = o(t)$, using just crude bounds on the third moments of $\Delta Z_1^*(t)$ and $\Delta Z_2^*(t)$.

As we shall indicate shortly, each of the second-order terms on the right side of (16) converges to a finite limit when we divide it by t and let $t \downarrow 0$. The first-order terms can be rewritten, using (14), as

$$f_1(z)\mu_1 t + f_2(z)\mu_2 t + [f_2(z) - f_1(z)]\, E_z^*[M_2^*(t)] .$$

Thus it follows from (7) that $[T_t^* f(z) - f(z)]/t$ converges to a finite limit iff $f_2(z) - f_1(z) = 0$. When this holds, we combine (16) with (14) and (8)-(11) to obtain

$$(17) \qquad [T_t^* f(z) - f(z)]/t \to D^* f(z) - \frac{1}{2}(\sigma_2^2 + \sigma_{12}^2)(f_{12} - f_{11})(z)$$

as $t \downarrow 0$. In this calculation, we have also used the obvious relationships

$$E_z^*[(\Delta X_1^*(t))^2]/t \to \sigma_1^2 \qquad \text{as } t \downarrow 0,$$

$$E_z^*[(\Delta X_1^*(t)\ X_2^*(t)]/t \to \sigma_{12}^2 \qquad \text{as } t \downarrow 0,$$

and

$$E_z^*[(X_2^*(t))^2]/t \to \sigma_2^2 \qquad \text{as } t \downarrow 0.$$

Since $(f_2 - f_1) = 0$ along the axis $z_2 = 0$, it follows that $(f_{21} - f_{11}) = 0$ along this axis, and hence the second term on the right in (17) is zero.

In order to complete the proof, we need only show that $[T_t^* f(z) - f(z)]/t$ is bounded as a function of t and $z \in S$ jointly, at least for sufficiently small t. It is easy to show that $E_z^*[(\Delta Z_1^*(t))^2]/t$, $E_z^*[\Delta Z_1^*(t)\ \Delta Z_2^*(t)]/t$, and $E_z^*[(\Delta Z_2^*(t))^2]/t$ are bounded jointly in z and t. Thus, from Taylor's Theorem and the definition of $\mathcal{C}$, it will suffice to show that

$$\{f_1(z)\ E_z^*[\Delta Z_1^*(t)] + f_2(z)\ E_z^*[\Delta Z_2^*(t)]\}/t$$

is bounded jointly in t and z. Finally, using the fact that

$$\Delta Z_1^*(t) = \Delta X_1^*(t) - Y_2^*(t) \qquad \text{and} \qquad \Delta Z_2^*(t) = \Delta X_2^*(t) + Y_2^*(t)$$

for $t \in [0,T^*]$, together with the boundedness of the second-order partial derivatives of f and the fact that $(f_2 - f_1)(z_1,0) = 0$, we find it suffices to show that

$$\{z_2 E_z^*[Y_2^*(t)]\}/t = \{z_2 E_z^*[(M^*(t) - z_2)^+]\}/t$$

is bounded jointly in t and $z = (z_1, z_2)$. But this is easy to verify from the known first passage and maximum distributions of (one-dimensional) Brownian Motion, so the proof is complete.

6. A First Passage Problem

We now show that $\Phi(z) = P_z^*\{T^* = \infty\}$. Thus, determining the limit distribution of Z is equivalent to solving a first passage problem for Z^*.

Theorem 1. $\Phi_t(z) = P_z^*\{T^* > t\}$ for $t \geq 0$ and $z \in S$.

Corollary. $\Phi(z) = P_z^*\{T^* = \infty\}$ for $z \in S$.

Proof. We first rewrite Proposition 3 as

$$(18) \qquad \Phi_t(z_1, z_2) = P_0\{z_1 - M_1(t) \geq 0,\ z_1 + z_2 - M_2(t) \geq 0\}.$$

Next observe that

$$(19) \qquad z_1 - M_1(t) = z_1 - \sup_{0 \leq u \leq t} \{X_1(u)\} = \inf_{0 \leq u \leq t} \{z_1 - X_1(u)\}.$$

$$(20)\qquad z_1 + z_2 - M_2(t) = z_1 + z_2 - \sup_{0 \le v \le u \le t} \{X_2(v) + X_1(u)\}$$

$$= \inf_{0 \le v \le u \le t} \{[z_2 - X_2(v)] + [z_1 - X_1(u)]\} .$$

Thus, defining

$$M_1^*(t) = \inf_{0 \le u \le t} \{X_1^*(u)\} , \qquad t \ge 0 ,$$

$$M_2^*(t) = \inf_{0 \le v \le u \le t} \{X_2^*(v) + X_1^*(u)\} , \qquad t \ge 0 ,$$

we combine (18)-(20) to obtain

$$(21)\qquad \Phi_t(z) = P_z^*\{M_1^*(t) \ge 0,\ M_2^*(t) \ge 0\} = P_z^*\{[M_1^*(t) \wedge M_2^*(t)] \ge 0\} .$$

Finally, observe that

$$M_1^*(t) \wedge M_2^*(t) = \inf_{0 \le u \le t} \{X_1^*(u) \wedge [X_1^*(u) + \inf_{0 \le v \le u} X_2^*(v)]\}$$

$$= \inf_{0 \le u \le t} \{X_1^*(u) - Y_2^*(u)\} = \inf_{0 \le u \le t} \{Z_1^*(u)\} .$$

Combining this with (21) and the definition of T^* completes the proof of the theorem. The corollary follows from the fact that $\Phi_t(z) \downarrow \Phi(z)$ as $t \uparrow \infty$. (See Section 4). This completes the proof.

Now let $\mathcal{C}_0$ be the set of continuous functions $f: S \to \mathbb{R}$ which satisfy the following two conditions.

(22) All partial derivatives of f exist and are continuous everywhere on S except (possibly) at the origin.

(23) Off of every open neighborhood containing the origin, f and all of its partial derivatives are bounded.

From (23) and the continuity of f it follows that f is bounded.

<u>Theorem 2</u>. Assume $\mu_1 < 0$ and $\mu_1 + \mu_2 < 0$. Suppose $\Psi \in \mathcal{C}_0$ satisfies

(24) $(\Psi_2 - \Psi_1)(z_1, 0) = 0$ for $z_1 > 0$,

(25) $D^*\Psi = 0$ except (possibly) at the origin,

(26) $\Psi(z_1, z_2) \to 0$ uniformly in z_2 as $z_1 \downarrow 0$,

(27) $\Psi(z_1, z_2) \to 1$ uniformly in z_2 as $z_1 \uparrow \infty$.

Then $\Psi = \Phi$.

<u>Remark 1</u>. There is strong reason to believe that the converse is also true, meaning that Φ necessarily satisfies (24)-(27). From the definition (6), it follows that Φ satisfies (26) and (27). It also follows from (6) that Φ satisfies (24) so long as both partial derivatives exist. Finally, from the corollary to Theorem 1 it follows easily that $\Phi \in \mathcal{D}^*$ and $G^*\Phi = 0$. To prove that Φ satisfies (25), it thus remains only to show that Φ belongs to a class of reasonably smooth functions (like $f \in \mathcal{C}_0$

such that D^*f is bounded) to which Proposition 6 $(G^*f = D^*f)$ can be extended.

Remark 2. It is important to note that Theorem 2 requires $\Psi \in \mathcal{C}_0$ but not $\Psi \in \mathcal{C}$. In Section 9 we present an example where Φ is in $\mathcal{C}_0$ and satisfies (24)-(27) but where all partial derivatives of Φ are unbounded in the neighborhood of the origin.

Proof. Suppose $0 < \epsilon < 1$, and assume $\epsilon < Z_1^*(0) < 1/\epsilon$. Let T_ϵ be the first time that Z_1^* hits either ϵ or $1/\epsilon$. Then T_ϵ is a stopping time for Z^*, and it is easy to show $E_z^*(T_\epsilon) < \infty$ for all z such that $\epsilon < z_1 < 1/\epsilon$. Proceeding exactly as in the proof of Proposition 6, one can shown that the function Ψ, restricted to the strip $\epsilon < z_1 < 1/\epsilon$, is in the domain of the generator of the process Z^*, modified by absorbing barriers at ϵ and $1/\epsilon$. Moreover, $G^*\Psi = D^*\Psi = 0$, where G^* now denotes the generator of the modified process. Applying Dynkin's formula, we then have

$$(28) \qquad E_z^*[\Psi(Z^*(T_\epsilon))] = \Psi(z) + E_z^* \int_0^{T_\epsilon} G^*\Psi(Z^*(t))\, dt = \Psi(z) \quad .$$

Observe that $T_\epsilon \to T^*$ almost surely as $\epsilon \downarrow 0$. Also, from (26) and (27) it follows that the left side of (28) converges to $P_z^*(T^* = \infty)$ as $\epsilon \downarrow 0$. Thus, letting $\epsilon \downarrow 0$ in (28) and using the Corollary to Theorem 1, we obtain $\Psi(z) = \Phi(z)$.

7. The Density Function

The remainder of the paper consists of detailed calculations for which subscripted variables are clumsy. Thus points in S will be denoted (x,y) hereafter, rather that (x_1,x_2) or (z_1,z_2). Let $\mathcal{C}^*$ be the set of functions $f:S \to \mathbb{R}$ that are continuous except (possibly) at the origin and that satisfy (22) and (23). A function in $\mathcal{C}^*$ may be unbounded in the neighborhood of the origin.

<u>Theorem 3</u>. Assume $\mu_1 < 0$ and $\mu_1 + \mu_2 < 0$. There exists a function $\Psi \in \mathcal{C}_0$ satisfying (24)-(27) iff there exists $f \in \mathcal{C}^*$ satisfying

(29) $$D^*f = 0 \text{ on the interior of } S\,,$$

(30) $$[\sigma_{12}^2 f_1 + \tfrac{1}{2}\sigma_2^2 f_2 - \mu_2 f](x,0) = 0 \qquad \text{for } x > 0,$$

(31) $$[\tfrac{1}{2}\sigma_1^2 f_1 + (\tfrac{1}{2}\sigma_1^2 + \sigma_{12}^2) f_2 - \mu_1 f](0,y) = 0 \qquad \text{for } y > 0,$$

(32) $$f \geq 0 \quad \text{and} \quad \int_0^\infty \int_0^\infty f(x,y)\, dx\, dy = 1\,.$$

In this case, $f = \Psi_{12} - \Psi_{22}$ and

(33) $$\Psi(z,y) = \int_0^x \int_0^{x+y-u} f(u,v)\, dv du \qquad \text{for } (x,y) \in S.$$

<u>Corollary</u>. If f satisfies (29)-(32), then f is a density for the limit distribution F.

Remark 1. The comments following Theorem 2, when combined with Theorem 3, suggest that (29)-(32) are both necessary and sufficient for the density of the limit distribution.

Remark 2. Again we emphasize that the theorem does not require $f \in \mathcal{C}$. In the example of Section 9, both f and its partial derivatives are unbounded in the neighborhood of the origin.

Proof. First suppose that $\Psi \in \mathcal{C}_0$ satisfies (29)-(32) and define $f = \Psi_{12} - \Psi_{22}$ except at the origin (where the value of f can be set arbitrarily). From the definition of $\mathcal{C}_0$ it follows that $f \in \mathcal{C}^*$. Also, since $(\Psi_1 - \Psi_2)(x,0) = 0$, we have

$$(34) \qquad \frac{\partial}{\partial x} \Psi(x,z-x) = (\Psi_1 - \Psi_2)(x,z-x) = \int_0^{z-x} f(x,v)\, dv$$

for $z > x > 0$. Next, since $\Psi(0,z) = 0$, we have

$$(35) \qquad \Psi(x,z-x) = \int_0^x \frac{\partial}{\partial u} \Psi(u,z-u)\, du = \int_0^x \int_0^{z-u} f(u,v)\, dv du$$

for $z > x > 0$, and (35) is equivalent to (33). Now from (35) we directly compute that $D^*\Psi = \Delta^* f$ on the interior of S, where

$$\Delta^* f(x,y) = \int_0^y D_1^* f(x,v) dv + \int_0^x D_2^* f(u,y) du + (\tfrac{1}{2}\sigma_1^2 + \sigma_{12}^2)\, f(x,y),$$

$$D_1^* f = \frac{1}{2}\sigma_1^2 f_1 - \mu_1 f \quad \text{on the interior of } S,$$

and

$$D_2^* f = (\frac{1}{2}\sigma_1^2 + \sigma_{12}^2 + \frac{1}{2}\sigma_2^2) f_2 - (\mu_1 + \mu_2) f \quad \text{on the interior of } S.$$

From the fact that $D^*\Psi = 0$ on the interior of S we have

$$(36) \qquad 0 = (\frac{\partial^2}{\partial x \partial y} - \frac{\partial^2}{\partial y^2}) D^*\Psi(x,y) = (\frac{\partial^2}{\partial x \partial y} - \frac{\partial^2}{\partial y^2}) \Delta^* f(x,y) ,$$

$$(37) \qquad 0 = (\frac{\partial}{\partial x} - \frac{\partial}{\partial y}) D^*\Psi(x,0) = (\frac{\partial}{\partial x} - \frac{\partial}{\partial y}) \Delta^* f(x,0) ,$$

$$(38) \qquad 0 = \frac{\partial}{\partial y} D^*\Psi(0,y) = \frac{\partial}{\partial y} \Delta^* f(0,y).$$

Directly computing the right-most expressions in (36)-(38) from (33), we obtain (29)-(31). Finally, Theorem 2 gives $\Psi = \Phi$, and then from (33) and the definition (6) of Φ it follows that f is a density for the proper distribution F. Thus f satisfies (32).

For the converse, suppose f satisfies (29)-(32) and define Ψ by (33). It is immediate from (32) and (33) that Ψ satisfies (24), (26) and (27). Thus we need only show $D^*\Psi = 0$. Again we have (36)-(38), the second equality in each case following from the fact that $D^*\Psi = \Delta^* f$ by (33), and the triple of first equalities being equivalent to (29)-(31). But (36)-(38) are necessary and sufficient for $D^*\Psi = K$ (a constant). Thus we need to show $K = 0$. Whatever the value of K, the proof of Theorem 2 need be altered only trivially to establish that

$$(39) \qquad E_z^*[\Psi(Z^*(T_\epsilon))] = \Psi(z) + KE_z^*(T_\epsilon).$$

When $\mu_1 < 0$ and $\mu_1 + \mu_2 < 0$, it is easy to show that $E_z^*(T_\epsilon) \to \infty$ as $\epsilon \downarrow 0$ whenever $z_1 > 0$. Since Ψ is bounded, it follows from (39) that $K = 0$.

8. Exponential Solutions

Since our basic differential equation $D^*f = 0$ has constant coefficients, it is natural to seek solutions of the separable exponential form

$$f(x,y) = \alpha\beta e^{-(\alpha x+\beta y)}, \qquad (x,y) \in S. \tag{40}$$

Proposition 7. Assume $\mu_1 < 0$ and $\mu_1 + \mu_2 < 0$. There exists a solution f of (29)-(32) having the form (40) iff $\frac{1}{2}\sigma_1^2 + \sigma_{12}^2 = 0$, in which case $\alpha = 2|\mu_1|/\sigma_1^2$ and $\beta = 2|\mu_1 + \mu_2|/\sigma_2^2$.

Remark. A solution f of (29)-(32) can have the form $f(x,y) = g(x)\,h(y)$ only if f is of the form (40). This follows easily from the boundary conditions (30) and (31). Thus the proposition implies that $\frac{1}{2}\sigma_1^2 + \sigma_{12}^2 = 0$ is necessary and sufficient for the existence of a separable solution.

Proof. Assuming the form (40), conditions (29)-(31) are equivalent to

$$\frac{1}{2}\sigma_1^2\alpha^2 + \sigma_{12}^2\alpha\beta + \frac{1}{2}\sigma_2^2\beta^2 + \mu_1\alpha + \mu_2\beta = 0\ . \tag{41}$$

(42) $$\sigma_{12}^2\alpha + \frac{1}{2}\sigma_2^2\beta + \mu_2 = 0 ,$$

and

(43) $$\frac{1}{2}\sigma_1^2\alpha + (\frac{1}{2}\sigma_1^2 + \sigma_{12}^2)\beta + \mu_1 = 0 .$$

Multiplying (42) through by β (which must be positive if f is to be a density) and subtracting this from (41), we obtain $\frac{1}{2}\sigma_1^2\alpha + \mu_1\alpha$. Comparing this with (43), we see that (41)-(43) can hold simultaneously iff $\frac{1}{2}\sigma_1^2 + \sigma_{12}^2 = 0$, in which case $\alpha = 2|\mu_1|/\sigma_1^2$. Substituting this value of α in (42) gives $\beta = 2|\mu_1 + \mu_2|/\sigma_2^2$.

9. Another Special Case

When the density f does not have the simple form (40), computing it explicitly appears to be a difficult problem. Only one other special case has been solved thus far.

Proposition 8. Assume $\sigma_{12}^2 = 0$, $\sigma_1^2 = \sigma_2^2 = 1$, $\mu_2 = 0$, and $\mu_1 < 0$. Let $\mu = |\mu_1|$. Then the density of the limit distribution F is

$$f(x,y) = Ke^{-\mu x}[r(x,y)]^{-1/2} e^{-\mu r(x,y)} \cos[\frac{1}{2}\theta(x,y)] ,$$

where $r(x,y) = (x^2 + y^2)^{1/2}$, $\theta(x,y) = \tan^{-1}(y/x)$ and $K = 2\pi^{-1/2}(2\mu)^{3/2}$.

Remark. The slightly more general case $\sigma_1^2 = \sigma_2^2 = \sigma^2$ can of course be accommodated by a rescaling.

Proof. In this case our equations (29) through (32) for f become

(44) $$\frac{1}{2} f_{11} + \frac{1}{2} f_{22} + \mu f_1 = 0 \qquad \text{on the interior,}$$

(45) $$\frac{1}{2} f_2(x,0) = 0 \qquad \text{for } x > 0$$

(46) $$[\frac{1}{2} f_1 + \frac{1}{2} f_2 + \mu f](0,y) = 0 \qquad \text{for } y > 0 .$$

Letting $g(x,y) = e^{\mu x} f(x,y)$, we re-express (44) through (46) in terms of g as

(47) $$g_{11} + g_{22} = \mu g \qquad \text{on the interior,}$$

(48) $$g_2(x,0) = 0 \qquad \text{for } x > 0 ,$$

(49) $$[g_1 + g_2 + \mu g](0,y) = 0 \qquad \text{for } y > 0 .$$

Finally, we transform to polar coordinates by letting $h(r,\theta) = g(x,y)$. The usual transformation formulas then give (47)-(49) in terms of h as

(50) $$r^{-2} h_{22}(r,\theta) + r^{-1} h_1(r,\theta) + h_{11}(r,\theta) = \mu^2 h(r,\theta)$$

(51) $$r^{-1} h_2(r,0) = 0 \qquad \text{for } r > 0 ,$$

(52) $$h_1(r, \pi/2) - r^{-1} h_2(r, \pi/2) + \mu h(r, \pi/2) = 0 \qquad \text{for } r > 0.$$

Hypothesizing a solution of the form $h(r,\theta) = \varphi(r)\,\psi(\theta)$, we find that $\varphi(r) = r^{-1/2}\exp(-\mu r)$ and $\psi(\theta) = \cos(\theta/2)$ satisfy (50)-(52). Reversing the transformations then gives the desired formula for f except that the normalization constant K has not been determined. We need

$$1 = \int_0^\infty \int_0^\infty f(x,y)\,dxdy = K \int_0^\infty \int_0^\infty [r(x,y)]^{-1/2} e^{-\mu[x+r(x,y)]} \cos[(\tfrac{1}{2}\theta(x,y)]dxdy$$

$$= K \int_0^{\pi/2} \int_0^\infty r^{-1/2} e^{-\mu[1+\cos\theta]r} \cos(\tfrac{1}{2}\theta) r\,drd\theta \ .$$

Let $\lambda(\theta) = \mu[1 + \cos\theta]$, so $\lambda(\theta) = 2\mu\cos^2(\frac{1}{2}\theta)$. Observe that

$$\int_0^\infty r^{1/2} e^{-\lambda(\theta)r}\,dr = \lambda^{-3/2}(\theta) \int_0^\infty u^{1/2} e^{-u}\,du = \Gamma(\tfrac{3}{2})\,\lambda^{-3/2}(\theta)$$

$$= \tfrac{1}{2}\sqrt{\pi}\,\lambda^{-3/2}(\theta) = \tfrac{1}{2}\sqrt{\pi}\,[2\mu\cos^2(\tfrac{1}{2}\theta)]^{-3/2}$$

$$= \tfrac{1}{2}\sqrt{\pi}\,(2\mu)^{-3/2}\cos^{-3}(\tfrac{1}{2}\theta)$$

Thus we need

$$1 = K[\tfrac{1}{2}\pi^{1/2}(2\mu)^{-3/2}] \int_0^{\pi/2} \cos^{-2}(\tfrac{1}{2}\theta)d\theta = K[\tfrac{1}{2}\pi^{1/2}(2\mu)^{-3/2}][2\tan(\tfrac{1}{2}\theta)]_0^{\pi/2}$$

Thus $K = 2\pi^{-1/2}(2\mu)^{3/2}$.

10. Concluding Remarks

Probably the best known work on multi-dimensional diffusions with boundaries is the seminal paper by Stroock and Varadhan [6]. Strictly speaking, our process Z lies outside the Stroock-Varadhan theory, because its boundary fails to meet their smoothness requirements. The same is true of the theory developed by Watanabe [7].

In Section 1 we have discussed only the fact that the equilibrium distribution of the diffusion Z approximates the distribution of αW for a tandem queuing system, where W is the equilibrium waiting time vector. One can, however, show that in heavy traffic the distribution of the entire (vector) waiting time process is approximated by the distribution of Z in the following sense. Assume a stable tandem queuing system ($\alpha > 0$), and define a two-dimensional process

$$\zeta(t) = \alpha W_{[t/\alpha^2]}, \qquad t \geq 0,$$

where $[\cdot]$ denotes integer part. Now consider a sequence of stable tandem systems with $\alpha \downarrow 0$. Under assumptions like those employed in our previous paper [2], and using very much the same methods, it can be shown that ζ converges weakly (in the appropriate function space) to Z as $\alpha \downarrow 0$. In a similar vein, one might examine the normalized (vector) queue length process

$$\nu(t) = \alpha Q(t/\alpha^2), \qquad t \geq 0,$$

where $Q(t) = [Q_1(t), Q_2(t)]$ and $Q_k(t)$ denotes the number of customers present at station k at time t. Iglehart and Whitt [3] have proved heavy traffic limit theorems for processes of this form associated with a general class of acyclic queuing networks. (Their normalizations, however, are expressed in a rather different form.) If one specializes their results for ν to the case of two queues in tandem, one finds that ν converges weakly to Z as $\alpha \downarrow 0$, where Z is essentially the same diffusion studied here.

The last paragraph suggests that we might be interested in quantities associated with Z other than its equilibrium distribution. In particular, the quantity $\Phi_t(\cdot)$ provides an approximation for the <u>transient</u> waiting time or queue length distribution of a tandem system, provided that the system starts empty and heavy traffic conditions prevail. Theorem 1 and Proposition 6 together suggest that $\Phi_t(\cdot)$ should satisfy

$$D^*\Phi_t(z) - \frac{\partial}{\partial t}\Phi_t(z) = 0 \quad \text{for } t > 0 \text{ and } z \text{ in the interior of } S,$$

together with the boundary conditions $\Phi_t(0, z_2) = 0$ and

$$\left(\frac{\partial}{\partial z_2} - \frac{\partial}{\partial z_1}\right)\Phi_t(z_1, 0) = 0 \quad \text{for } t > 0 \text{ and } z_1 > 0 .$$

and the initial condition $\Phi_0(\cdot) \equiv 1$. We shall make no attempt to prove this, except for the following comment. If a sufficiently smooth solution can be found, then Dynkin's formula can be used to prove that this solution is $\Phi_t(\cdot)$, just as in the proof of Theorem 2.

Acknowledgment. I am indebted to Professor S. Varadhan for his generous assistance during the last year of this research. In particular, the illuminating example discussed in Section 9 would not have been solved without his stimulus.

References

[1] Dynkin, E.B. (1965), Markov Processes, Vol. 1, Springer-Verlag, New York.

[2] Harrison, J.M. (1973), The Heavy Traffic Approximation for Single Server Queues in Series, J. Appl. Prob., Vol. 10, 613-629.

[3] Iglehart, D.L. and W. Whitt (1970), Multiple Channel Queues in Heavy Traffic. II: Sequences, Networks, and Batches, J. Appl. Prob. Vol. 2, 355-369.

[4] Itô, K. (1961), Lecture Notes on Stochastic Processes, Tata Institute for Fundamental Research, Bombay.

[5] Reiman, M. (forthcoming), Diffusion Approximations for Queuing Networks in Heavy Traffic, Ph.D. Dissertation, Dept. of Operations Research, Stanford University.

[6] Stroock, D.W., and S.R.S. Varadhan (1971), Diffusion Processes with Boundary Conditions, Comm. Pure Appl. Math., Vol. 24, 147-225.

[7] Watanabe, S. (1971), On Stochastic Differential Equations for Multi-dimensional Diffusion Processes with Boundary Conditions, J. Math. Kyoto Univ., Vol. 11, 164-180.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

14

1. REPORT NUMBER: TR-75, TR-45

2. GOVT ACCESSION NO.

3. RECIPIENT'S CATALOG NUMBER: 9

4. TITLE (and Subtitle): 6 THE DIFFUSION APPROXIMATION FOR TANDEM QUEUES IN HEAVY TRAFFIC.

5. TYPE OF REPORT & PERIOD COVERED: TECHNICAL REPORT,

6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s): 10 J. MICHAEL HARRISON

8. CONTRACT OR GRANT NUMBER(s): 15 N00014-76-C-0418, NSF-ENG-75-14847

9. PERFORMING ORGANIZATION NAME AND ADDRESS:
DEPT. OF OPERATIONS RESEARCH
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: NR-047-061

11. CONTROLLING OFFICE NAME AND ADDRESS:
Logistics & Mathematical Statistics Branch
Code 434 Office of Naval Research
Arlington, Va. 22217

12. REPORT DATE: AUGUST 10, 1977

13. NUMBER OF PAGES: 33

14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office): 11 10 Aug 77 12 37p.

15. SECURITY CLASS. (of this report): UNCLASSIFIED

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE, DISTRIBUTION UNLIMITED.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

This research was supported in part by National Science Foundation Grant ENG 75-14847 Department of Operations Research, Stanford University and issued as Technical Report No. 45

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

DIFFUSION PROCESS, TANDEM QUEUE, QUEUING NETWORK, REFLECTING BARRIER, HEAVY TRAFFIC, MARKOV PROCESS

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

SEE REVERSE SIDE

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

402 766

THE DIFFUSION APPROXIMATION FOR TANDEM QUEUES IN HEAVY TRAFFIC

J. Michael Harrison*

Graduate School of Business
Stanford University

ABSTRACT

Consider a pair of single server queues arranged in series. (This is the simplest example of a queuing network.) In an earlier paper [2], a limit theorem was proved to justify a heavy traffic approximation for the (two-dimensional) equilibrium waiting time distribution. Specifically the waiting time distribution was shown to be approximated by the limit distribution F of a certain vector stochastic process Z. The process Z was defined as an explicit, but relatively complicated, transformation of vector Brownian Motion, and the general problem of determining F was left unsolved.

It is shown ~~in this paper~~ that Z is a diffusion process (continuous strong Markov process) whose state space S is the non-negative quadrant. On the interior of S, the process behaves as an ordinary vector Brownian Motion, and it reflects instantaneously at each boundary surface (axis). At one axis, the reflection is normal, but at the other axis it has a tangential component as well. The generator of Z is calculated.

It is shown that the limit distribution F is the solution of a first passage problem for a certain dual diffusion process Z*. The generator of Z* is calculated, and the analytical theory of Markov process is used to derive a partial differential equation (with boundary conditions) for the density f of F. Necessary and sufficient conditions are found for f to be separable (for the limit distribution to have independent components). This extends slightly the class of explicit solutions found previously in [2]. Another special case is solved explicitly, showing that the density is not in general separable.

75/45